Syllabus for the course

How to examine and predict time series: methods and applications

[Time series prediction and signal processing]

How to communicate with a computer by voice? How to forecast rates of a national currency? How to identify an oncoming heart attack in good time? – the answers to these questions are associated with the problems (1) to recognize and synthesize speech, (2) to predict a chaotic time series, and (3) to reveal typical sequences in an observed time series, respectively. All these problems along with many others comprise the field of time series prediction.

The course combines real-world applications with a strong theoretical background: the authors selected mathematical topics required to solve complex problems of actual practice. On the other hand, several topics focus on "mathematics of future" that is theories that will have become the basis of applications in the decades to come. The course starts with simple concepts and gradually works in more advanced applications.

To be specific, the course deals with main models to examine and predict regular time series (exemplified by ARIMA and GARCH models), chaotic time series (predictive clustering, constructive neural networks, deep learning models and others) as well as with state-of-the-art approaches used to distinguish regular and chaotic time series, using observations of the time series at hand only; particular topics deal with the concepts of forecasting (time) and stationarity horizons and methods of signal processing. Applications considered range from econometrics problem over speech recognition and synthesis to mobile health.

The assessment includes 2 intermediate tasks and final exam.

Cumulative score = 0.5*Score for the 1st intermediate task + 0.5*Score for the 2nd intermediate task.

Final score = 0.6*Cumulative score + 0.4*Final exam score.

The final exam is oral. The prerequisite for the course is a course in basic statistics.

TOPIC 1. **Correlation analysis**. Causal connection. The Pearson's product-moment coefficient. Its properties. Statistical tests for Pearson's coefficient. The partial correlation coefficient. The coefficient of multiple correlation. Statistical tests for it.

TOPIC 2. Linear regression analysis. Theoretical and applied regression definitions. Gauss-Markov conditions. The least squares method. Properties of least squares estimators. BLUE. Remainder variance. Tests for parameters and line of regression. How to compare regression models? How to compare estimated parameters corresponding to different samples? The weighted least squares.

TOPIC 3. Non-linear regression analysis. The correlation ration. Its properties. Non-linear regression analysis. Linearized regression and its properties.

TOPIC 4. **Neural networks and time series predictions.** Constructive neural networks. NEAT-model. Deep learning and time series prediction.

TOPIC 5. Data balancing techniques. Rationale for data-balancing. Shift algorithm with modified remainders. Reparametrized models. Iterative algorithm to adjust parameters.

TOPIC 6. Predictive regression models. Unconditional prediction. Conditional prediction. Autoregression models prediction.

TOPIC 7. Maximum likelihood estimation of regression parameters. Properties of maximum likelihood estimation. Maximum likelihood estimation for linear models. Statistical tests. Non-linear constraints.

TOPIC 8. Regular time series. Time series. Unit roots and co-integration. Autocorrelation and partial autocorrelation functions. Properties of AR(1)- and AR(2)- processes. Properties of moving average processes.

TOPIC 9. ARIMA-models. Identification of ARIMA-models. Method to estimate model parameters. Forecasting with ARIMA. GARCH-models.

TOPIC 10. **Non-linear and chaotic time series**. **Forecasting horizon**. Chaos “fingerprints”. Forecasting horizon. How to calculate forecasting horizon for a given time series? Multiplicative ergodic (Oseledets) theorem. An invariant measure of a dynamic system. Kolmogorov-Sinai entropy: series that generate information.

TOPIC 11. **Attractor reconstruction for chaotic time series.** Takens theory. Optimal reconstruction parameters. Properties of correlation integral. Limitations of non-linear dynamics algorithms. Econophysics.

TOPIC 12. Predictive clustering. One and multi-step ahead prediction. Clustering algorithms employed for predictive clustering. Association with invariant measure. Non-predictable observations. Similar time series. Relation tensor of time series.

TOPIC 13. Applications of predictive clustering. Mobile health. A plant to produce technical indicators for stock markets. Prediction of hash-tag popularity. Weather forecasting. Energy consumption prediction. Text of literature pieces as chaotic time series.

TOPIC 14. Signal processing. Discrete signals. Basic definitions. Convolution. The Dirac delta function and FIR. Convolution with respect to input and output. Properties of convolution.

TOPIC 15. Discrete Fourier transform. The real-input DFT. Basic functions. Signal synthesis with DFT. Signal analysis with DFT. Calculation of DFT. Applications of DFT. Its properties.

TOPIC 16. Introduction to digital filtration. Basic definitions. Representation information on signal. Time and frequency characteristics. Types of filters. Homogenous filter. Applications.

TOPIC 17. Audio signal processing. Human hearing. Sound timbre. Tradeoff between sound quality and discretization frequency. Speech synthesis and recognition. Non-linear audio signal processing.

TOPIC 18. Video signal processing. Brightness and contrast. Linear video signal processing. Convolution. High-dimensional DFT. Fast Fourier transform.

TOPIC 19. Stationarity horizon. Non-stationary time series. Definition of stationarity horizon. Empirical probability density function. How to compare two

PDFs? Estimated minimum sample size. Horizon series. Its probability density function.

TOPIC 20. Bifurcation precursors and time series. The concept of bifurcation precursors (early-warning signs). Various approaches to identify pre-bifurcation states with the employment of time series.

TOPIC 21. Predicitve complexity. Predictability and learning. Learning a parameterized process. Beyond finite parameterization: general considerations and model process.

1. Unpingco J. Python for Signal Processing. Switzerland: Springer International Publishing, 2014.
2. Prandoni P., Vetterli M. Signal processing for communications. Boca Raton, FL: Taylor and Francis Group, 2008.
3. Fan J., Yao Q. Nonlinear Time Series: Nonparametric and Parametric methods, N.-Y., Springer, 2003.
4. Malinetskii G.G., Potapov A.B. Modern problems of non-linear dynamics. Moscow: URSS, 2002. (in Russian)
5. Magnus Ya.R., Katyshev P.K., Pereseckii A.A. Introduction to econometrics. Moscow: Delo, 2007. (in Russian)
6. Orlov Yu.N., Osminin K.P. Nonstationary time series: prediction methods and applications to financial and raw materials markets. Moscow: URSS, 2018. (in Russian)